Subgroups of New York City Children at High Risk of PTSD After the September 11 Attacks: A Signal Detection Analysis

Craig S. Rosen, Ph.D.
Michael Cohen, Ph.D.

Objective: **Case finding is an important challenge in mental health programs responding to large-scale disasters. Most people who experience psychological symptoms after such events return to normal functioning within a few months. Yet a significant minority continues to experience enduring symptoms. This study demonstrated the use of signal detection analyses of community survey data to identify subgroups of children who were at highest risk of posttraumatic stress disorder (PTSD) after the September 11 attacks.** ***Methods:*** **This study reanalyzed results of a needs assessment survey conducted six months after the World Trade Center attacks on September 11, 2001, with a representative sample of 7,832 New York City public school students in grades 4 to 12. Receiver operating characteristic (ROC) analyses conducted on half the sample resulted in a decision tree for classifying children into groups at varying levels of risk of PTSD. These decision rules were subsequently retested on the second half of the sample.** ***Results:*** **We could reliably classify children into groups with varying probabilities of screening positive on a PTSD screen. Nearly two-thirds of children in grades 4 to 12 who screened positive for probable PTSD were concentrated among 4th graders (35%) and among children who had a friend or family member directly exposed to the attacks (28%).** ***Conclusions:*** **Signal detection analysis of community needs assessment surveys can identify community subgroups most likely to screen positive for mental health problems after a disaster or terrorist attack. This information can help target screening and outreach efforts to community segments that have the highest need for services. (*Psychiatric Services* 61:64–69, 2010)**

Case finding is a significant challenge in delivering mental health services after natural disasters, industrial disasters, or terrorist attacks (1). Most people who experience acute distress after a disaster return to normal functioning within six months to a year after the event (1). However, a significant minority of those who are distressed continue to have enduring symptoms and problems months or years after the event. Children and teens are especially likely to have emotional difficulties after disasters (2). However, children and adults who have mental health problems after a disaster often go undetected, undiagnosed, and untreated (3).

Disaster mental health programs in the United States typically use proactive outreach efforts to try to penetrate all segments of the affected community (4). Such universal outreach is somewhat inefficient, because most people who receive outreach contacts would recover spontaneously without any mental health services (1). Outreach to the entire affected community may be a reasonable case-finding strategy in localized disasters that have an impact on a circumscribed community: before 2001, the median federally funded crisis counseling grant in the United States served 1,500 clients (4).

Universal outreach is a less effective case-finding strategy when responding to events that have an impact on millions of people. It is virtually impossible to individually contact and screen everyone who might need mental health services after events such as Hurricane Katrina, the 2004 tsunami disaster, or the September 11, 2001, attacks. When responding to such massive events, it is critical that finite mental health outreach, screening, and treatment resources be targeted to where they can do the most good.

What data could inform such targeting decisions? In recent years there have been several efforts using epidemiological surveys with generalizable sampling procedures to accurately assess communities' mental health needs after disasters. Many of these studies have focused on children (3,5–7). Such community surveys are often conducted several months after the event, when many survivors have already experienced remission of their acute symptoms,

Dr. Rosen is affiliated with the National Center for PTSD, Department of Veterans Affairs Palo Alto Health Care System, 795 Willow Rd. (334-PTSD), Menlo Park, CA 94025 (e-mail: craig.rosen@va.gov). He is also with the Department of Psychiatry and Behavioral Science, Stanford School of Medicine, Stanford, California. Dr. Cohen is president of the Michael Cohen Group, L.L.C., New York City.

and can help determine the need for continuing services after the immediate postdisaster phase.

The efficiency of case finding might be improved by focusing screening and outreach efforts on community segments that have the highest prevalence of problems in a community survey. However, the logistic regression techniques typically used to analyze epidemiological surveys do not generate simple decision rules that can be used to assign individuals to different risk groups. An alternative approach is using receiver operating characteristic (ROC) analysis, a signal detection method commonly used in medical decision making to classify individuals into groups that are at higher or lower risk of having a medical condition or an adverse health outcome (8,9). ROC analyses generate a series of decision rules for sorting individuals into higher- and lower-risk groups.

The study presented here examined the potential value of ROC analyses of community survey data in identifying groups of children who are at the highest risk of posttraumatic stress disorder (PTSD) after a disaster. This study reanalyzed data from a survey that assessed mental health needs of New York City public school children six months after the September 11 attacks (10,11). The original needs assessment study determined both overall mental health needs and identified numerous risk factors associated with mental health problems. However, the original study was not designed to generate decision rules for differentiating higher- and lower-risk subgroups within the sample.

This reanalysis used ROC methods on the first half of the sample to generate a set of decision rules for identifying subgroups of children who were more or less likely to score positive on a PTSD screen. The resultant decision rules were then retested in the second half of the sample. The study addressed two questions related to reliability and detection: How well did risk segments developed in the ROC analysis replicate in the second half of the sample? How many high-risk children (those who screened positive for probable PTSD) were included in the various risk groups?

Some readers may be familiar with using ROC analysis to assess the validity of a screening tool against a gold-standard diagnostic criterion. Our purpose was different. We used ROC analysis to identify risk factors associated with children screening positive on a PTSD screen. We were not assessing the validity of the screening measure itself, which has been validated in prior research.

Methods

Sample

The original needs assessment study (10,11) was commissioned by the New York City Board of Education (now the New York City Department of Education). The study was conceived and conducted by Applied Research and Consulting, L.L.C. (ARC), with significant contribution from researchers at the Columbia University Mailman School of Public Health. The sampling design was created by researchers at ARC and the Centers for Disease Control and Prevention. The survey was conducted with a representative sample of 8,236 students in grades 4 to 12, of whom 7,832 students (95%) completed the PTSD screening measure. The sample was designed to represent the entire population of New York City public school students in grades 4 to 12, with oversampling of several subpopulations. Subsamples that were oversampled included students from schools in the immediate physical vicinity of Ground Zero, students from schools that had a clear view of the World Trade Center, and students who went to schools in areas that had a high proportion of public safety workers residing in the community. Details of the clustered sampling procedures have been published previously (11).

Several factors make this a particularly appropriate survey for testing the potential value of ROC in case finding. First, it was explicitly designed as a needs assessment study intended to inform service delivery. Second, the study used well-designed sampling procedures that enable projection of results to the entire population of New York City public school students who were in grades 4 to 12. Third, the sample was large enough to enable split-half replication of the ROC model.

Measures

Probable mental health disorders were assessed with the Diagnostic Interview Schedule for Children Predictive Scales (DPS) (12). The original study measured eight different disorders, but most disorders were assessed in only a subsample of participants (planned missingness). Because ROC analysis does not allow for imputation of missing data, we decided to focus our analyses on one disorder that was assessed among all the participants. We selected PTSD because it was assessed specifically in reference to the World Trade Center attacks and because it was the psychiatric disorder with the most elevated prevalence (fivefold increase) relative to comparison data from other community samples (10). In prior unpublished research cited by Hoven and colleagues (11), the PTSD screen had 85% sensitivity and 98.4% specificity.

A questionnaire on the September 11, 2001, World Trade Center attacks was developed specifically for the study and assessed various exposure factors, including physical proximity (for example, attending a school near Ground Zero) and direct exposure to stressors (such as witnessing the attack, being near the cloud of smoke, being evacuated from one's school, and having difficulty returning home after the attack), psychological proximity (whether the child knew anyone killed in the attacks or anyone who survived the attacks and his or her relationship to that person), and amount of exposure to media coverage of the attack. The three psychological proximity variables were recoded as ordinal scales: 0, don't know anyone; 1, acquaintance ("someone else you know"); 2, friend; 3, nonimmediate family member; 4, sibling; and 5, parent.

The survey also assessed children's self-report of exposure to traumas before the September 11 attacks, and these did predict children's risk of PTSD. We intentionally handicapped ourselves by not including prior exposure data in the reanalysis presented here, because we assumed that men-

tal health staff trying to plan a disaster response would typically not have access to such detailed information of individual students' histories.

Procedures

The 50 data collectors completed human subjects training and were trained in administration of the study protocol. Parental consent information was distributed to parents via students' schools. Fourth and fifth graders were included in the study if the parents returned a consent form (active consent); older children were included unless their parents opted out of the study (passive consent). The measures were completed in school classrooms within one class period. Fourth and fifth graders filled in their responses as the questions were read aloud to them. Older children completed paper-and-pencil surveys on their own. A 24-hour telephone number for mental health assistance and referral was provided to all families in the study. (Because the surveys were truly anonymous, children could not be referred on the basis of their screening responses.) The study was overseen by the institutional review boards of the New York City Department of Education, the New York State Office of Mental Health Committee for WTC-Related Research, and Columbia University–New York State Psychiatric Institute.

Analyses

Data analysis proceeded in three phases. The first was an ROC analysis conducted on the first half of the sample (N=3,945). This analysis identified risk factors that classified children into groups according to whether they were more or less likely to screen positive for probable PTSD (probable PTSD was defined as above-threshold symptoms accompanied by self-reported impairment in functioning). Each predictor variable was examined to determine its single best cutoff point for differentiating participants into high-risk and low-risk groups. We then selected the one variable that best split the sample into high- and low-risk groups. This analysis was iteratively repeated within each resultant risk group until additional predictors became nonsignificant or the cell sizes became too small (fewer than ten cases). This resulted in an empirically derived decision tree for classifying individuals into higher- and lower-risk groups.

In the second phase, we confirmed in the second of half of the sample (N=3,887) whether the ROC-generated decision tree reliably determined children's likelihood of screening positive for PTSD. We assigned children to groups using the decision tree developed in the first half of the sample. We then conducted two sets of significance tests, both using the z test for differences in proportions (13). The first set of tests, conducted only in the replication half of the sample, compared the two groups generated at each point of the decision tree to see whether each branching point again generated groups with differing rates of positive PTSD screens. For this analysis, finding a significant difference across the groups would confirm the potential utility of the decision rule. The second set of tests compared results across both samples and tested whether children in each risk category had similar rates of positive PTSD screens in both the model development and replication samples. For these tests, finding no significant difference across the two samples would confirm the reliability of the model.

In the third analysis, we combined both halves of the sample and reweighted it to reflect the entire population of grades 4 through 12 New York City public school students. This enabled us to better estimate the proportion of all students and the proportion of students likely to screen positive for probable PTSD that would be included in each group identified in the ROC model.

Results

ROC model generation

Results of the ROC decision tree are shown in Figure 1. Two ratios are shown within each box. The top ratio is from the model development half of the sample; the ratio on the bottom in italics is from the replication half of the sample. In all the ratios, the denominator is the total number of children in each subgroup of the sample; the numerator is the number of children in that subgroup who screened positive for probable PTSD; the percentage is the proportion of children in that segment who screened positive for probable PTSD.

Grade level emerged as the single strongest predictor of probable PTSD. The ROC identified the optimal split-point as being the distinction between grade 4 (27% screening positive for probable PTSD) and grades 5 to 12 (7% screening positive). Among children in grade 4, those who had been directly exposed to two or more stressors were more likely than those who had not been to screen positive for PTSD (39% versus 17%). Among fourth graders who were exposed to less than two stressors, those who knew anyone who was at the World Trade Center during the attack were more likely than those who did not to screen positive (27% versus 11%).

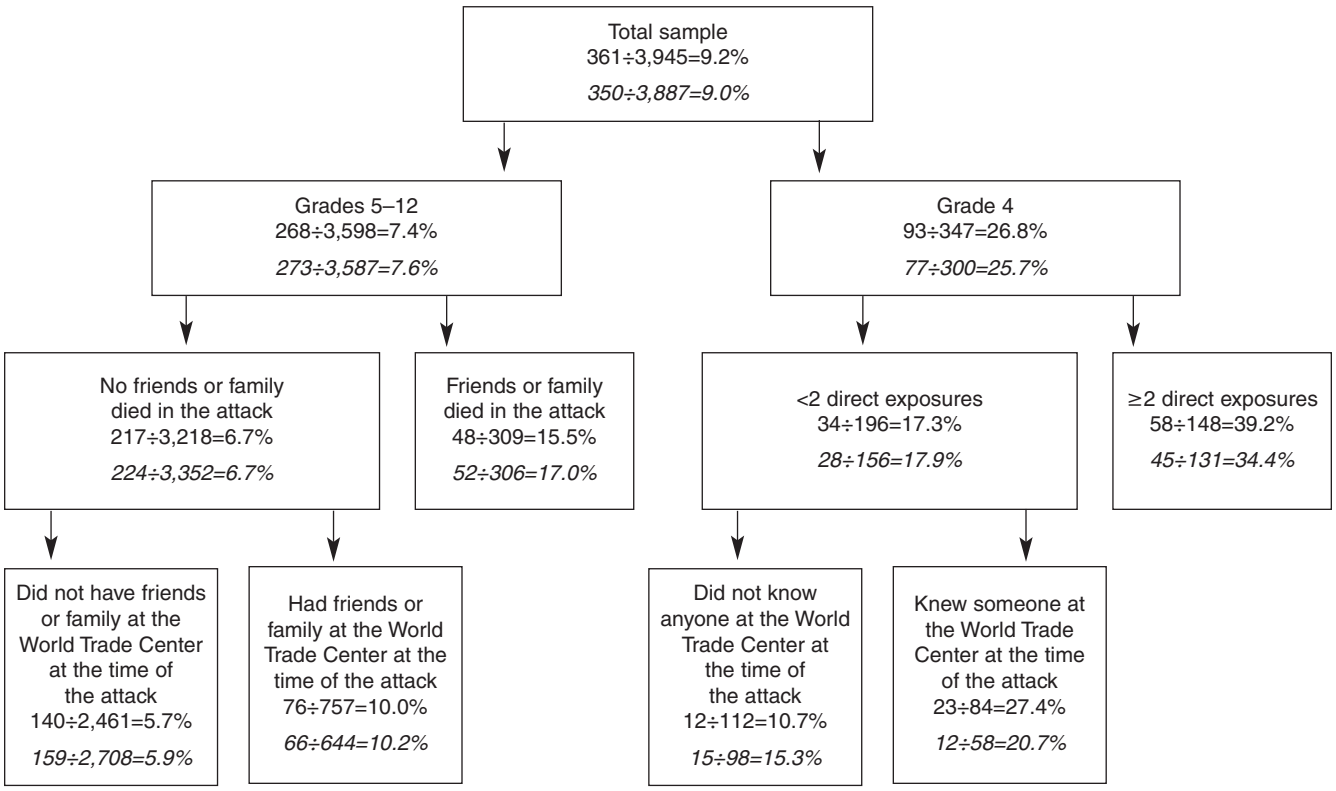
Among children in grades 5 to 12, the strongest predictor of screening positive for probable PTSD was whether a friend or family member died in the attack. Sixteen percent of those whose friends or family members died in the attack had probable PTSD, whereas those whose friends and family did not die in the attack were not as likely to screen positive (7%). The next highest predictor in these grade levels was having a friend or family member who was at the World Trade Center at the time of the attack and survived (10%, screened positive). Fifth to twelfth graders who did not have a friend or family member at the World Trade Center at the time of the attack were less likely to screen positive for PTSD (6%).

Replication

We next applied the ROC-generated decision rules to the second (replication) half of the sample. These results are shown in the second set of ratios in Figure 1. Our first set of z tests for proportions, conducted only within the replication sample, tested whether each split in the decision tree yielded groups with differing rates of positive PTSD screens. We found significant differences ($z=3.18$ to 10.49, $p<.01$ to $p<.001$) for all but one

Figure 1

Likelihood of New York City public school students screening positive for probable posttraumatic stress disorder, as identified by the receiver operating characteristic (ROC) model[a]



[a] Top numbers are results from the first (model development) half of the sample. Bottom numbers (italics) are from the second (replication) half of the sample.

branch of the tree, confirming the potential utility of these splits. The one split that did not replicate was among fourth graders exposed to fewer than two direct exposures: among these children, the rate of positive PTSD screens no longer differed significantly depending on whether or not the child knew someone who had been at the World Trade Center at the time of the attacks. This distinction was dropped from our final model.

Our second set of z tests for proportions tested whether the rates of positive PTSD screens within each of the risk categories were similar in both the model development and replication subsamples. These were all nonsignificant, confirming that the model was reliable across samples.

Projection to the fourth- to 12th-grade student population

Having confirmed that the ROC-generated segmentation rules reliably differentiated children's risk of screening positive for probable PTSD, we projected the results to the population of New York City public school students in grades 4 to 12. Results are shown in Table 1 for each of the five risk segments.

The group most likely to screen positive for PTSD was fourth graders exposed to multiple trauma cues. We estimated that this group constituted only about 5% of the grade 4 to 12 public school student population yet nearly one-fifth of all children who would have screened positive for probable PTSD. Fourth graders who experienced fewer than two direct exposures were estimated to be 9% of the population and accounted for roughly 16% of children likely to screen positive on the PTSD measure. We estimated that children in grades 5 to 12 who lost a friend or family member in the attacks comprised roughly 7% of the population and about 12% of all children who would have screened positive for probable PTSD. Fifth to 12th graders who had a friend or family member who had been at the World Trade Center at the time of the attack and survived accounted for roughly 17% of both the study population and of children who would have screened positive for PTSD. Finally, we estimated that fifth to 12th graders who had no friends or family at the World Trade Center at the time of the attacks would have included over 63% of the total study population yet only 36% of all children who would have screened positive for probable PTSD. Combining the three highest-risk groups (two or more exposures and fewer than two exposures for fourth graders; family or friend died in the attack for fifth to 12th graders), we estimated that nearly half (47%) of the children who would have screened positive on the PTSD screen were concentrated in roughly one-fifth of the grade 4 to 12 public school student population.

Table 1

Estimated distribution of posttraumatic stress disorder (PTSD) cases in grades 4 to 12 in the New York City public school system, by risk category[a]

Risk category	Projected total population (N=7,816)[b]		Projected population who would screen positive for PTSD (N=706)	
	N	%	N	%
Grade 4				
≥2 exposures	279	5.3	103	19.2
<2 exposures	352	8.6	62	16.3
Grade 5–12				
Family or friend died	615	6.6	100	11.8
Family or friend survived the World Trade Center attacks	1,401	16.7	142	16.5
No family or friend at the World Trade Center during the attacks	5,169	62.8	299	36.2

[a] Ns are unweighted, and percentages are weighted to reflect the entire population of the New York City public school system (grades 4 to 12).

[b] N is 7,816 because 16 cases had missing values for direct exposure or exposure of friends or family.

Discussion

This study demonstrated the potential utility of using ROC analyses of needs assessment survey data to develop risk strata that could inform targeting of mental health outreach and screening after a large-scale disaster or terrorist attack. Our findings identified two different types of high-risk groups. One risk pathway involved having a family member or friend survive or be killed in the attacks. A sensible service delivery strategy to reach these children would be to screen children who knew people directly exposed to the attacks (and especially children who lost someone), conduct follow-up clinical assessments with those who screened positive for PTSD, and provide counseling services to those who needed them.

A second risk group was defined by grade: fourth graders were much more likely than older children to screen positive for PTSD. This suggests three (not mutually exclusive) possibilities: effects of age-related differences in the method of administration, the DPS PTSD screen's being less specific (generating more false positives) among fourth graders than among children in later grades, or higher rates of pathology among the fourth graders. We can rule out the differences in administration. Both fourth and fifth graders had the questions read aloud to them and participated only if their parents gave active consent. Yet the proportion of positive PTSD screens among fourth graders (27%) was double that of fifth graders (N=70 of 504, or 14%) who were assessed with the same procedures.

It could be that the PTSD screen is less specific (generates more false positives) among fourth graders than among older children. Fourth graders are at the extreme bottom of the nine- to 17-year-old age range used to develop the DPS, and its developers did not report its psychometrics for different age subgroups (12). Yet other indicators suggest that there may in fact have been greater rates of PTSD among younger children: fourth graders were most likely to endorse PTSD if they had been exposed to two or more stressors. Fourth graders also reported more PTSD and anxiety disorders than did older children but less depression and conduct disorder, suggesting that if there was any overreporting bias, it was disorder specific, not universal.

A sensible next step in service delivery would be to administer both a PTSD screen and a comprehensive clinical assessment to a sample of elementary school students and older children. This would determine the rates of false positives generated by the PTSD screen and assess whether they varied by age. If the elevated prevalence of PTSD among younger children appeared to be genuine, more services could be directed to elementary grades, especially in schools that were evacuated, exposed to lingering smoke, or exposed to other stressors.

The risk factors for PTSD identified in this study (greater exposure to disaster-related stressors, death of friends or loved ones, and younger age) are not novel. They are consistent with the broader disaster research literature (1). The key innovation of the methods used here is the ability to translate risk factors into a set of decision rules that parsimoniously and reliably differentiate children into groups at higher and lower risk of PTSD. Had these analyses been available and been used to inform targeting of outreach and screening efforts, we might have been able to deploy services more efficiently to reach more of the children who most needed help.

At the same time, the study confirms that an empirically based case-finding strategy is unlikely to identify all of the children likely to need services. In our study, roughly one-third of all children screening positive for PTSD were not in any of the high-risk groups. Moreover, not all children who have PTSD would be detected by a brief epidemiological screen. This suggests that targeted screening efforts must also be accompanied by mechanisms for referring, assessing, and treating any child who appears to be highly distressed, regardless of his or her risk group.

A useful feature of ROC is that it generates risk groups with homogeneous risk factors, unlike logistic regression, which uses weighted averages of different risk factors (9). This makes the resulting segments intuitive and potentially useful for tailoring interventions. For example, our study suggested that classroom-based screening might have been helpful in reaching fourth graders, because entire classrooms were in a high-risk category. In contrast, older children would have best been detected with individual screening and referral of children who lost friends or loved ones.

Another strength of ROC analyses is that the resultant decision trees are easily interpretable by policy makers and constituents (14). Decisions about allocating resources after disasters are politically sensitive; tensions often erupt over perceived (or actual) inequities in the distribution of recovery assistance (4). Having a clear and face-valid rationale can help policy makers defend their decisions in targeting mental health services.

Our study has several limitations. As noted above, additional research is needed to confirm whether the DPS PTSD screen is accurate across the entire developmental range for grades 4 to 12. Our needs assessment survey did not include children in kindergarten to grade 3, who were assumed to lack the needed literacy skills to respond to the measure. Better methods are needed to efficiently screen such young children, because other research suggests that young children may be at particularly high risk after disasters (1,2). Another limitation was that the original survey was designed with planned missingness so that some questions were given to only a subsample of respondents. ROC analyses have no provision for imputing such missing data, limiting the number of disorders that we could include in our analysis.

A final limitation is that community survey data are often not available after most disasters. Such surveys tend to be conducted only after particularly large or newsworthy disasters or attacks. However, it is primarily for responding to large-scale events that better targeting information is most needed.

Conclusions

One challenge in providing mental health services after disasters is efficiently directing services to the minority of people who are not already recovering on their own. After large-scale disasters or terrorist attacks, it is virtually impossible to screen the entire affected community. Signal detection analyses of community surveys can reliably identify community subgroups that are the most likely to screen positive for mental health problems. Such risk profiles can be used to efficiently target postdisaster outreach and screening efforts to specific community groups that are likely to have the highest need for mental health services.

Acknowledgments and disclosures

This reanalysis was supported by the Department of Veterans Affairs (VA) National Center for PTSD and the VA Palo Alto Health Care System. The views expressed here are those of the authors and do not necessarily reflect the position of the U.S. Department of Veterans Affairs. The original needs assessment survey was funded by an award from the New York City Department of Education (NYC DOE) to Applied Research and Consulting, L.L.C. (Dr. Cohen, contract investigator). The original study was a collaboration among Applied Research and Consulting, the Mailman School of Public Health, Columbia University-New York State Psychiatric Institute (Cristina Hoven, Dr.P.H., principal investigator), and the NYC DOE. Special thanks go to our collaborators at Turnaround for Children, Inc. (Pamela Cantor, M.D.); Columbia University-New York Psychiatric Institute (Cristina Hoven, Dr.P.H., Cristiane Duarte, Ph.D., M.P.H., Christopher P. Lucas, M.D., Ping Wu, Ph.D., Donald Mandell, Ph.D., M.P.H., Renee Goodwin, Ph.D., M.P.H., Bin Fan, M.D., George Musa, M.S., Larry Aber, Ph.D., Patricia Cohen, Ph.D., Ezra Susser, M.D., Dr.P.H., Judith Wicks, B.A., Andrea Versenyi, B.A., Barbara P. Aaron, B.A., Henian Chen, M.D., Mark Davies, M.P.H., Steven Greenwald, Ph.D., and Patricia Zybert, Ph.D.); Applied Research and Consulting (Nellie Gregorian, M.P.A., Chris Bumcrot, B.A., and Victoria Francis, M.A.); the Centers for Disease Control and Prevention (Bradley Woodruff, M.D., M.P.H., and Victor Balaban, Ph.D.); the National Center for Child Traumatic Stress, University of California, Los Angeles (Robert Pynoos, M.D., M.P.H., Alan Steinberg, Ph.D., and William Saltzman, Ph.D.); and Steven Marans, Ph.D., Elissa Brown, Ph.D., Claude Chemtob, Ph.D., and Betty Pfefferbaum, M.D. This study could not have succeeded without the leadership of Francine Goldstein, M.A., and the participation of Vincent Giordano, M.S., Linda Wernikoff, M.A., all from the NYC DOE, as well as the superintendents, principals, teachers, and, most of all, the students from New York City.

The authors report no competing interests.

References

1. Norris FH, Friedman MJ, Watson PJ, et al: 60,000 disaster victims speak: part I. an empirical review of the empirical literature, 1981–2001. Psychiatry 65:207–239, 2002
2. Somasundaram DJ, van de Put WA: Management of trauma in special populations after a disaster. Journal of Clinical Psychiatry 67:64–73, 2006
3. Fairbrother G, Stuber J, Galea S, et al: Unmet need for counseling in New York City after the September 11 attacks on the World Trade Center: implications for pediatricians. Pediatrics 113:1357–1366, 2004
4. Norris FH, Rosen CS, Elrod, CL, et al: Retrospective 5-Year Evaluation of the Crisis Counseling Program. White River Junction, Vt, Veterans Affairs National Center for PTSD, 2005
5. Pfefferbaum B, Nixon SJ, Krug RS, et al: Clinical needs assessment of middle and high school students following the 1995 Oklahoma City bombing. American Journal of Psychiatry 156:1069–1074, 1999
6. Roussos A, Goenjian AK, Steinberg AM, et al: Posttraumatic stress and depressive reactions among children and adolescents after the 1999 earthquake in Ano Liosia, Greece. American Journal of Psychiatry 162:530–537, 2005
7. Thienkrua W, Cardozo BL, Chakkraband ML, et al: Symptoms of posttraumatic stress disorder and depression among children in tsunami-affected areas in southern Thailand. JAMA 296:549–559, 2006
8. Balemans WA, van der Ent CK, Schilder AG, et al: Prediction of asthma in young adults using childhood characteristics: development of a prediction rule. Journal of Clinical Epidemiology 59:1207–1212, 2006
9. Kiernan M, Kraemer HC, Winkleby MA, et al: Do logistic regression and signal detection identify different subgroups at risk? Implications for the design of tailored interventions. Psychological Methods 6:35–48, 2001
10. Hoven CW, Duarte CS, Cohen MC, et al: Effects of the World Trade Center Attack on NYC Public School Students. New York, New York City Board of Education, 2002
11. Hoven CW, Duarte CS, Lucas CP, et al: Psychopathology among New York City public school children 6 months after September 11. Archives of General Psychiatry 62:545–552, 2005
12. Lucas CP, Zhang H, Fisher PW, et al: The DISC Predictive Scale (DPS). Journal of the American Academy of Child and Adolescent Psychiatry 40:443–449, 2001
13. Fleiss JL: Statistical Methods for Rates and Proportions, 2nd ed. New York, Wiley, 1981
14. Rosen CS, Young HE, Norris, FH: On a road paved with good intentions, you still need a compass: monitoring and evaluating disaster mental health services; in Interventions Following Mass Violence and Disasters. Edited by Ritchie EC, Watson PJ, Friedman MJ. New York, Guilford, 2006